D4.4

85-08

Vandewalle J., Vanderschoot J., De Moor B., ``Filtering of vector signals based on the singular value decomposition'', in *7th European Conference on Circuit Theory, Design*, Prague, 2-6 Sep. 1985, pp. 458-461., Lirias number: 181841.

FILTERING OF VECTOR SIGNALS BASED ON THE SINGULAR VALUE DECOMPOSITION.

J. Vandewalle)[+], J. Vanderschoot)[++] and B. De Moor)[+ +++]

)[+]ESAT Laboratory, Electrical Engineering Department
Katholieke Universiteit Leuven
3030 Heverlee, Belgium
Tel. 016/22.09.31

)[++]Laboratory of G.I. Motility
Department of Medical Research
Katholieke Universiteit Leuven
3030 Heverlee, Belgium

Abstract: Adaptive and partial singular value decomposition of sets of samples of vector signals allows to perform filtering and signal separation. This technique is applicable to all cases where many signals composed of many sources (signals of different strength and noise) can be measured simultaneously. It has been applied to the extraction of the fetal ECG out of many cutaneous measurements containing maternal, fetal ECG and noise.

1. INTRODUCTION

In this paper we present an alternative way of signal filtering based on geometry along with its algorithmic implementation and practical use. If a set of measured signals contains contributions from sources with different strengths, they can often be separated by projecting onto subspaces of maximal and minimal energy. This technique has been applied successfully to filter or separate the fetal electrocardiogram (ECG) from the maternal ECG in cutaneous measurements, and to filter out 50 Hz contributions, and to eliminate noise in the state space realization.

A numerically reliable algorithm to perform this separation requires the singular value decomposition SVD of the matrix of the measured signals. Since both the classical SVD algorithm of Golub and an adaptive SVD algorithm ASVD are too complex for real time processing, a new and faster adaptive algorithm is developped.

In the paper we first define the oriented energy of a vector signal and derive a number of its properties [1]. Then the signal separation property is derived and applied to the separation of the fetal ECG from cutaneous measurements [2] . Third a numerically reliable signal separation algorithm based on SVD is presented. A new adaptive algorithm based on the Chebyshev transformation is then presented which is much faster than the classical algorithm of Golub [3] and the adaptive SVD algorithm ASVD [4] .

2. ORIENTED ENERGY IN A SIGNAL VECTOR SPACE

Suppose that A_{mxn} is a matrix, representing a set of n vector samples a_i from an m dimensional signal space, with n>m :

$$A = [a_1 \; a_2 \; \ldots\ldots \; a_n] \qquad (1)$$

+++Supported by the Belgian I.W.O.N.L.

Then the energy of this set in the direction of a unit vector e can be defined as [1]:

$$E_e[A] = \sum_{i=1}^{n} (e^T a_i)^2 \qquad (2)$$

It is the sum of the squared lengths of the orthogonal projections of all a_i onto e. The oriented energy in a subspace is equivalently, the sum of the squared lengths of the orthogonal projections of all a_i onto that subspace.

For example Fig. 1 gives polar plots of the oriented energy of sets of vectors in a three dimensional signal space (m=3). In general, for higher dimensions, polar plots of oriented energy all have the same properties: orthogonal directions of extremal energy and a maximum, a minimum, saddlepoints.

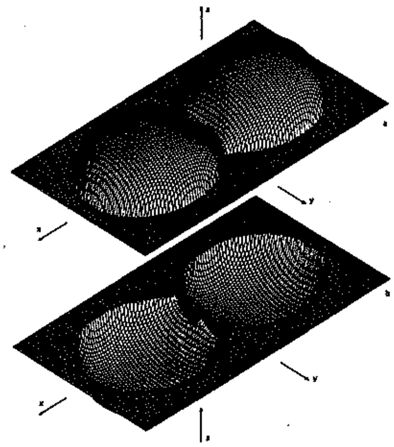

Fig. 1: Plot of the oriented energy of a 3 dimensional signal.

Although these orientations can be computed from the quadratic form of the autocorrelation matrix AA^t, it is well known [1,3] that this squaring often causes numerical problems. These can be avoided by using the singular value decomposition, which is recently receiving much attention in signal processing and system theory.

It is shown in linear algebra, that every real matrix A_{mxn} can be written as a product :

$$A_{mxn} = U_{mxn} \Sigma_{mxn} N_{nxn}$$

in which U and V are orthonormal matrices, and Σ is a real pseudodiagonal nonnegative definite matrix. The singular spectrum of A is the set of diagonal elements ("singular values") of Σ. The number of non zero singular values is equal to the rank of matrix A. The left singular basis is the set of column vectors of U. Conventionally the singular values are ordered in decreasing order of magnitude along the diagonal of Σ. The singular spectrum gives a reliable tool for the rank estimation of a matrix A, which is corrupted by inaccuracies. If the rank r of this noisy matrix is less than m (with $m \leq n$), the last m-r singular values are not zero. If the inaccuracies are random, these m-r singular values all have about the same magnitude. Their magnitudes are also smaller than those of the first r singular values.

The algebraic properties of the SVD of A, relating it to the oriented energy, can be easily proven, and understood by geometrical intuition [2] :

1. The energy in the direction of a left singular vector is equal to the square of the corresponding singular value :

$$E_{u_i}[A] = \sigma_i^2 \qquad (4)$$

2. The p dimensional subspace S^p of maximal energy is spanned by the first p column vectors of U, and its maximal energy equals the sum of the p first squared singular values. So, the direction of maximal energy of A is given by u_1, and the maximal energy of A is equal to σ_1^2. Also the plane of maximal energy of A is spanned by (u_1,u_2) and its maximal energy is equal to $\sigma_1^2 + \sigma_2^2$.
3. The p dimensional subspace S^p of minimum energy is spanned by the p last column vectors of U, and its minimal projected energy equals the sum of the p last squared singular values. So, the direction of minimal energy of A is given by u_m, and its minimal energy is equal to σ_m^2. If the matrix A is of rank r<m, then $\sigma_m = 0$. The direction of minimal non zero energy is then given by u_r.

Combining 2. and 3. we can say that the direction of maximal energy in the (m-1) dimensional subspace orthogonal on u_1 is given by u_2, etc.

3. DIRECTIONS OF EXTREMAL ENERGY AND SOURCE SEPARATION

We are interested in separating sources of different strength or extracting some sources from e measurement signals which are linear combinations of the instantaneous values of the sources corrupted by noise (Fig. 2). The technique is based on the following Theorem.

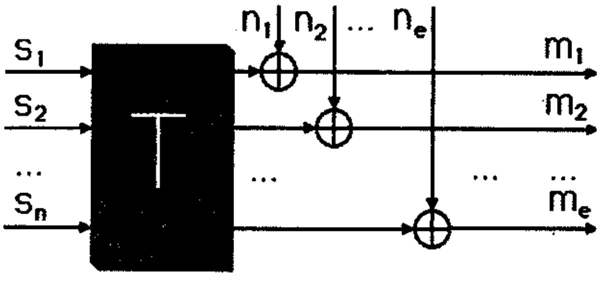

Fig. 2. The measured signals $m_1, m_2, \dots m_e$ are linear combinations of unknown sources $s_1, s_2 \dots s_n$ corrupted by noise $n_1, n_2 \dots n_e$.

THEOREM

Given a measurement matrix M :

$$M_{ext} = T_{exn} S_{nxt} + N_{ext}$$

with e the number of measurement signals,
t the number of vector samples,
n the number of source signals,
t>e>n
T,S,N matrices reperesenting respectively transfer, source signals, noise signals.

Let the singular value decomposition of M be given by :

$$M_{ext} = U_{exn} \Sigma_{nxn} V_{nxt}$$

If (1) $S S^T = \text{diag}[\sigma_{s_1} \ \sigma_{s_2} \ \dots\dots \ \sigma_{s_n}]$
implying that source signals are not correlated with each other for zero shift in time ,
(2) $\sigma_{s_1} > \sigma_{s_2} > \dots\dots\dots\dots > \sigma_{s_n} > 0$
(3) $T^T T = I$
implying that columns of T form an orthonormal set of basis vectors.
(4) $N N^T = \sigma_n^2 I$
implying that noise signals are not correlated, and that they all have the same energy level,
(5) $N S^T = 0_{exn}$
implying that noise signals are not correlated with the source signals.

Then (1) $\Sigma \Sigma^T = S S^T + \sigma_n^2 I$
(2) $U = T$
(3) $S = U^T M$

The proof of this theorem is based on the uniqueness of the eigenvalue decomposition of the matrix $M M^T$, and will be omitted here. Statement (3) of the theorem gives a solution to the inverse problem. To obtain this result, conditions (1) to (5) must be, at least approximately be satisfied. Conditions (2), (4) and (5) are quite easy to fulfil. On the contrary, it is not obvious that conditions (1) and (3) are easy to satisfy. Instead of a thorough discussion ot this matter, we extract the signal separation algorithm out of this theorem.

SIGNAL SEPARATION ALGORITHM

1. Compute the SVD (6) of a block M of t samples the measured data.
2. Compute the projection $\hat{s}_i \stackrel{\Delta}{=} u_i^T M$ of the measured sinals onto the directions of extremal energy $u_1, u_2 \dots u_e$ which are the columns of U. If the condition of the theorem are approximaly satisfied this is a good approximation of the source signal $s_i \simeq u_i^T M \stackrel{\Delta}{=} \hat{s}_i$

Of course one can in practice often reduce considerably the computations. Sometimes one is only interested in some source signals, then only the corresponding parts of the singular value decomposition have to be computed (partial singular decomposition). The singular value decomposition between two consecutive blocks may only slightly differ, then an adaptive SVD algorithm will be more effective. These computational aspects are discussed in the next section.

In this section, for a better understanding, an application of the method is described with a simulation example although the technique has been successfully applied on measured signals. This example concerns the measurement of a fetal vector cardiogram with electrodes located on the skin (abdominal FECG). Consider for simplicity two sources each with two signals : the maternal heart with $s_{mx}(t)$ and $s_{mz}(t)$, and the fetal heart with $s_{fx}(t)$ and $s_{fz}(t)$. Fig. 3. a represents a 4 sec. interval (1000 samples) of the four source signals.

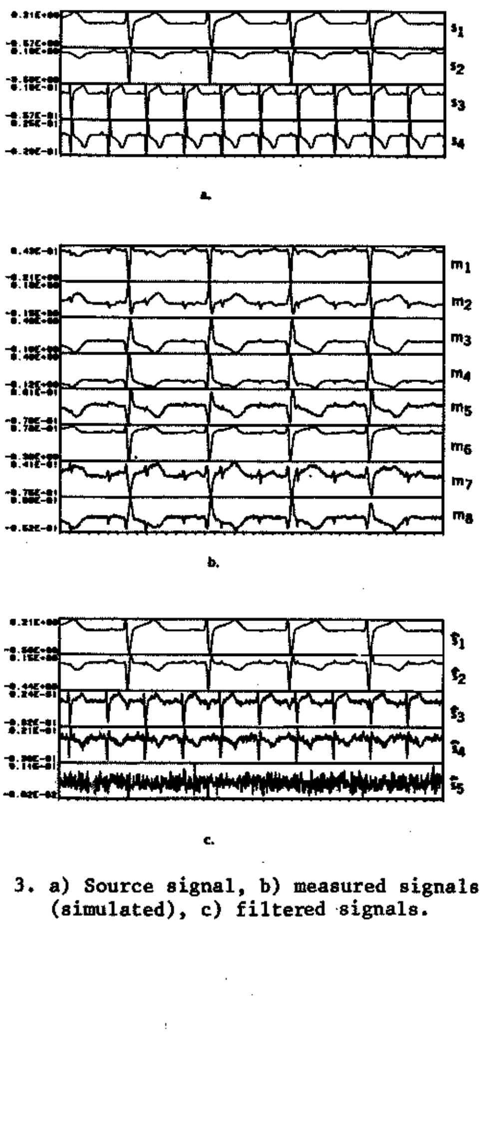

Fig. 3. a) Source signal, b) measured signals (simulated), c) filtered signals.

Fig. 3.b depicts 8 simulated measurement signals, obtained by making arbitrary linear combinations of the source signals. Some white Gaussian noise is also added to the electrode signals. The form of these measurement signals is quite similar to real recording. Fig. 3.c gives the projections of all 1000 measurement vector samples onto the directions of extremal energy based on the first 250 samples resulting in $\hat{s}_1(t)$ to $\hat{s}_5(t)$.

It is clear that, whereas in the measurement signals only the highest peaks of the fetal cardiogram can be seen, $\hat{s}_3(t)$ and $\hat{s}_4(t)$ allow the observation of much smaller features, and are practically free of maternal ECG. An additional benefit is that these projections on the orientations of extremal energy are independent of the physical orientation of the heart, which is of considerable practical importance. Observe that in this simulation example no special care has been taken to satisfy the conditions of the Theorem.

4. THE COMPUTATION OF THE PARTIAL OR ADAPTIVE SINGULAR VALUE DECOMPOSITION.

In this section we discuss some numerically reliable algorithms for the SVD of real matrices. The choice of a suitable algorithm for the computation of the SVD depends upon the conditions that are imposed. If one needs a full decomposition (i.e.all singular triplets) of a matrix of moderate dimensions, Golub's algorithm [3] is very efficient and numerically reliable. Fully tested and documentated software is available. However, there exist many applications in which the matrices involved satisfy specific conditions: They may be large and structured. such that the entries preferably remain unchanged during computations. When they are of low rank only a Partial SVD is necessary. (e.q. model reduction techniques). In other applications only the largest (l_2-norm of a matrix), the smallest (total LLS [6]) or some intermediate singular value (s) [2] are needed. A supplementary aspect may be that the matrices may be slowly time-varying. Hence, it is natural to develop an adaptive algorithm that uses the SVD computations of a previous time step as an initial value for updating the actual SVD. Both observations hold for many signal separation problems such as FECG. Two recently developped algorithms are now briefly considered.

A. ASVD : The power method for SVD

This method, developped independently by Staar [1] and Shlien [7] , and worked out in [5] can be sumarized as follows :

Suppose that k first triplets of an mxn matrix A are already computed

$$U_k = [u_1 \;\dots\dots\; u_k] \qquad (7a)$$

$$V_k = [v_1 \;\dots\dots\; v_k] \qquad (7b)$$

$$\Sigma_k = \mathrm{diag}\,[\sigma_1 \;\dots\; \sigma_k] \qquad (7c)$$

then the k+1-th triplet is computed via an iterative process

1. Initial guess $u_{k+1}^{(o)} \quad u_k$

iteration i

2. $v_{k+1}^{(2i+1)} = A^t.u_{k+1}^{2i} / ||A^t u_{k+1}^{2i}||$ (8)
3. $\sigma_{k+1}^{2i+1} = ||A^t.u_{k+1}^{2i}||$ (9)
4. $u_{k+1}^{2i+2} = A.v_{k+1}^{2i+1} / ||Av_{k+1}^{2i+1}||$ (10)
5. $\sigma_{k+1}^{2i+2} = ||Av_{k+1}^{2i+1}||$ (11)

6. If the test for convergence

$$||u_{k+1}^{2i+2} - u_k^{2i}|| < \varepsilon \quad (12)$$

is satisfied start the recusion for triplet k+2 (deflate),
if not , go to 2.

In [5] it is proven that the number of multiplications n_p for triplet p is

$$n_p = (-\log \varepsilon_m - \log(g_p/g_{p+1}) \; / \log (\sigma_p/\sigma_{p+1}) \quad (13)$$

where ε_m is the machine precision and g_p/g_{p+1} measures the quality of the initial guess.
Also convergence rates, reductions in the orthogonalization, initial guess strategies, deflation techniques and acceleration algorithms are studied in [5].

B. CHEBYCHEV ITERATION

This is a very efficient technique based on the recursion of the Chebychev polynomials.
From

$$A = U.\Sigma.V^t$$

one obtains

$$A.A^t = U.\Sigma.\Sigma^t.V^t \quad (14)$$

One can prove that for every polynomial or analytic function f of a matrix AA^t

$$f(AA^t) = V \;\; f(\Sigma\Sigma^t) \; V^t \quad (15a)$$

where $$f(\Sigma\Sigma^t) = \text{diag} \; [f \; (\sigma_i^2)] \quad (15b)$$

Now the Chebyshev polynomials $T_n(x)$ satisfy

$$T_o(x) = 1 \quad (16a)$$

$$T_1(x) = x \quad (16b)$$

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x) \quad (16c)$$

and $$|T_n(x)| > |T_{n-1}(x)| > 1 \quad x\varepsilon(-\infty,-1)\cup(1,\infty) \quad (17a)$$

$$|T_n(x)| < 1 \quad x\varepsilon \; [-1,1] \quad (17b)$$

If we now consider the following iterative procedure, with v_n the n- th iteration vector

1. Initialize v_o

$$v_1 = f \; (AA^t)v_o \quad (18)$$

2. Iterate $$v_{n+1} = 2f(AA^t)v_n - v_{n-1} \quad (19)$$

3. Test for convergence

$$||v_{n+1} - v_n|| < \varepsilon \quad (20)$$

if not, go to 2.

The result of this iterative procedure is a vector with components along those left singular vectors of A of which the corresponding singular value of $f\;(AA^t)$ is larger than 1. The largest, smallest or intermediary singular values are computed by choosing suitable f.
In [8], it is proven that the number of interations equels

$$n = -(\log \varepsilon_m + \log(g_p/g_{p+1})) \; /\log \; (d_p + \sqrt{d_p^2 - 1}) \quad (21)$$

where $d_p = f(\sigma_p^2)$

It can be proven that this iteration is generically considerably faster than the ordinary power method and Golub's algorithm [3,4]. Reliable refinements, deflation techniques and combination with previously developped acceleration techniques are actually under study.

5. CONCLUSION

A signal filtering technique has been described which is not based on the frequency content but rather on the strength of contribution to multiple signals. It has proven to be useful in eliminating maternal ECG in abdominal recordings of the fetal ECG. An algorithm for adaptive singular value decomposition which can compute efficiently a part of the SVD on-line

REFERENCES

[1] J. Staar, "Concepts for reliable modelling of linear systems with application to on-line identification of multivariable state space description", Doct. Th. , Kath. Univ. Leuven, June 1982.

[2] J. Vanderschoot, J. Vandewalle, J. Janssens, W.Sansen en G. Vantrappen, "Extraction of weak bioelectrical signals by means of singular value decomposition", "Analysis and optimization of systems", Proc. of Sixth Int. Conf. on An. and Opt. of Systems, Nice, June 1984, Springer Verlag, Berlin, pp. 434-448.

[3] G.H. Golub and C. Reinsch, "Singular value decomposition and least squares solutions", Numer. Math., Vol. 14, 1970, pp. 403-420.

[4] J. Dongarra et al. "Linpack User's Guide," Philadelphia, SIAM 1979.

[5] J. Vandewalle, J. Staar, B. de Moor and J. Lauwers, "An adaptive singular value decomposition algorithm and its application to adaptive realization", "Analysis and optimization of systems", Proc. of Sixth Int. Conf. on An. and Opt. of Systems, Nice, June 1984, Springer Verlag, Berlin, pp. 33-47.

[6] S. Van Huffel, J. Vandewalle and J. Staar, "The total linear least squares problem: formulation, algorithm and applications ", Proc. IEEE ISCAS, Montreal 1984, pp.328-331.

[7] S. Shlien, "A method for computing the partial SVD", IEEE Trans. on pattern analysis and mach. intell." Vol. PAMI-4, no 6, Nov. 1982.

[8] B. De Moor and J. Vandewalle, "An adaptive singular value decomposition algorithm based on generalized Chebyshev recursions", Submitted to Conf. on Math. in Signal Processing, Univ. of Bath, Essex, England, Sept. 85.